

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2002, 00:12:12 ; Search time 1693.08 seconds

(without alignments)
11838.176 Million cell updates/sec

Title: US-09-088-435-2

Perfect score: 1485
Sequence: 1 GGGTCGACACCGTCGCATG.....TTCTGATTTGAAAAAAA 1485

Scoring table:

IDENTITY_NIC
Gapop 10.0, Gapect 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

EST:
1: em_estb1:
2: em_estb2:
3: em_estb3:
4: em_estb4:
5: em_estb5:
6: em_estb6:
7: em_estb7:
8: em_estb8:
9: em_estb9:
10: em_estb10:
11: em_estb11:
12: em_estb12:
13: em_estb13:
14: em_estb14:
15: em_estb15:
16: em_estb16:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	773	52.1	773	10	BG679339
2	748	50.4	845	10	B1858271
3	744.2	50.1	815	10	BG702609
4	736.4	49.6	948	10	BG749122
5	726.4	48.9	807	10	BG473546
6	712.2	48.0	1099	10	BM463961
7	677	45.6	732	10	BF056112
8	666	44.8	667	10	BG774012
9	663.4	44.7	716	10	BG912245
10	646	43.5	679	9	A1953998
11	638	43.0	666	9	A1741818
12	605	40.7	606	9	BE222929
13	603	40.6	645	10	BE273896
14	587.2	39.5	964	10	BG169688
15	583.6	39.3	614	9	AA573827
16	563	37.9	700	10	B1670014
17	561.6	37.8	577	9	A1742027

18	517.8	34.9	570	10	BM129226
19	506.6	34.1	574	10	BM129278
20	502.6	33.8	807	10	BG768991
21	498.2	33.5	856	10	BE785282
22	492.2	33.1	806	10	BG968751
23	482.4	32.5	897	10	BG437057
24	480	32.3	480	10	BF064129
25	459.6	30.9	824	11	AK014487
26	449.4	30.3	451	9	A1344703
27	447.4	30.1	450	9	A1687918
28	435.2	29.3	600	10	BF780159
29	431.2	29.0	436	9	AM613258
30	427.2	28.8	591	10	B1359570
31	425.6	28.6	560	9	AM654331
32	425	28.6	602	9	BE134859
33	423.6	28.5	559	9	AM654338
34	417	28.1	418	9	A1241423
35	406.8	27.4	501	10	BE619177
36	400.8	27.0	573	10	B1682987
37	394.6	26.6	555	9	AM658374
38	394	26.5	546	9	AM656343
39	392	26.4	551	9	AM654557
40	392	26.4	565	10	B1541815
41	387.6	26.1	545	9	AM655435
42	386	26.0	562	9	AM656964
43	384.6	25.9	700	10	BE963955
44	378	25.5	542	9	AM656338
45	369.6	24.9	448	10	BF064050

ALIGNMENTS

RESULT 1
LOCUS BG679339 773 bp mRNA linear EST 01-MAY-2001
DEFINITION 602628516F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4753394 5',
mRNA sequence.
ACCESSION BG679339.1 GI:13910723
VERSION BG679339.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 773)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strusberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM10612 row: p column: 03
High quality sequence starters: 772.
Location/Qualifiers
1. 773
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4753394"
/clone.lib="NCI_CGAP_Skn4"
/tissue.type="squamous cell carcinoma"
/lab.host="DH10B (FLI phase-resistant)"
/note="Organ: skin; Vector: pCMV-SpOrf6; Site_1: NCI;
Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT

142 a 208 c 279 g 144 t

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaapps-remail.nih.gov

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC).
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCM/D/DP
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LNL)
Clone distribution by: Incyte Genomics, Inc.
Found through the I.M.A.G.E. Consortium/LNL at:
<http://Image.lnl.gov>
Plate: LAM12001 row: e column: 10
High quality sequence set: 819.
Location/Qualifiers
1. 845

```

1. 161 a
    236 c
    266 g
    162 h
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:5392861"
    /clone_1fp="NH_MGC_87"
    /tissue_type="mammary adenocarcinoma, cell line"
    /lab_host="DH10B (phage-resistant)"
    /note="Organ: breast; Vector: pCMV-Sport6; Site:1: NotI;
    Average insert size: 1.385 kb. Library enriched for
    full-length clones and constructed by Life Technologies.
    Note: this is a NIH_MGC Library."

```

[illegible]

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

Db 580 CCCCCCTCCCTGACTGATGTGTGTATCTGTGTGGAGGCAACATCATACAGCCGA 639

OY 1242 GAGCTGAGGCTTTGAAACCCACCTGGGCGAGGCTGAGGGGTCCCTCTGG-CCCCA 1300

Db 640 GAGCTGAGGCTTTGAAACCCACCTGGGCGAGGCTGAGGGGTCCCTCTGG-CCCCA 699

OY 1301 AAGACCCCTGAGAGGCCCATGAGACGTCCTGTCTGTGATGAGAGAGGAGTCTGCTGGC 1360

Db 700 AAGACCCCTGAGAGGCCCATGAGACGTCCTGTCTGTGATGAGAGAGGAGTCTGCTGGC 759

OY 1361 AGATGGAGAGTGG-AGAGTGGCCTGTGCACTGTGCTGCTCTCTTAAGAGAGCCCTC 1419

Db 760 AGATGGAGAGTGGAAAGGCTGCTGCAATGTGTGCTGCTCTCTTAAGAGAGCCCTC 819

OY 1420 C-TGAGCTGCTCTTTTGGCTCTC 1442

Db 820 CTTGGACTGGTCTTTTGGCTCTC 843

RESULT 3

LOCUS BG702609 815 bp mRNA linear EST 07-MAY-2001

DEFINITION 602684322F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4817021.5,

ACCESSION BG702609

VERSION BG702609.1 GI:13974113

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS 1 (bases 1 to 815)

TITLE NIH-MGC http://mgi.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: c9apds-remail.nih.gov

Tissue Procurement: Miklos Palokovits, M.D., Ph.D.

CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov

plate: LMA10716 row: k column: 06

High quality sequence stop: 811.

Location/Qualifiers

1. 815

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4817021"

/clone_id="NIH_MGC_95"

/tissue_type="hippocampus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site_1: BamHI, Site_2: SalI-XhoI (gtcagag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.5 kb and normalized to R0F 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH-MGC Library."

BASE COUNT 149 a 216 c 300 g 149 t 1 others

ORIGIN

Query Match 50.1%; Score 744.2; DB 10; Length 815;

Best Local Similarity 98.2%; Pred. NO. 1.7e-158;

Matches 805; Conservative 0; Mismatches 9; Indels 6; Gaps 5;

OY 201 AGCTGTTCCCTGGAGGCTGGAGGCGAGGAGCTGGGCCAGGCTGCTGGATAGTTG 260

Db 1 AGCGGTTCCCTGGAGGCTGGAGGCGAGGAGCTGGGCCAGCGTGTGGATAGTTG 60

OY 261 GCGAGGAGGCTGTGTACCTGTCTCCAGATGACGGCTGTGGCAGAGCATGCCAAG 320

Db 61 GCGAGGAGGCTGTGTACCTGTCTCCAGATGACGGCTGTGGCAGAGCATGCCAAG 120

OY 321 CAGGAGCCCTTTCACGTGGTCACTCTGTGTGAGAGCTGGGCGCTGCCAGGTGGG 380

Db 121 CAGGAGCCCTTTCACGTGGTCACTCTGTGTGAGAGCTGGGCGCTGCCAGGTGGG 180

OY 381 GCGATGCTGTCTTCTCAAGTGTGAGATGTGACAGCCAGCGCTCTTCAAGATTGG 440

Db 181 GCGATGCTGTCTTCTCAAGTGTGAGATGTGACAGCCAGCGCTCTTCAAGATTGG 240

OY 441 GCGATTGGGCAATTTCTGCCAGAGATGGCCAAAGAGGATGACACCTGTGTCTCC 500

Db 241 GCGATTGGGCAATTTCTGCCAGAGATGGCCAAAGAGGATGACACCTGTGTCTCC 300

OY 501 TCAGGGGGTAATGGCGGATCGCTGCTATGCTGCTAGGAGCTGGGCAATTCCTGCC 560

Db 301 TCAGGGGGTAATGGCGGATCGCTGCTATGCTGCTAGGAGCTGGGCAATTCCTGCC 360

OY 561 ACCATGCTGTCTCCCGAGAGCACTCCCTGACAGTGTGTGACAGGCTGCGAGGGAGGG 620

Db 361 ACCATGCTGTCTCCCGAGAGCACTCCCTGACAGTGTGTGACAGGCTGCGAGGGAGGG 420

OY 621 GCCGAGGTTACACTACTGTGAAAGTCTGTGGACAGAGCCAAATCTGAGGGCGCAAGATTG 680

Db 421 GCCGAGGTTACACTACTGTGAAAGTCTGTGGACAGAGCCAAATCTGAGGGCGCAAGATTG 480

OY 681 GCCAGAGGAGGAGGCTGGGAGATGCTCCCGGTTGACACCCCTAATAGGAAAGG 740

Db 481 GCCAGAGGAGGAGGCTGGGAGATGCTCCCGGTTGACACCCCTAATAGGAAAGG 539

OY 741 CACGCGAGCTGTGTGTCAGAGAGCTGAAAGCATGTGTGAGAGACCCCAAGAGTCCCTGTGTG 800

Db 540 CACGCGAGCTGTGTGTCAGAGAGCTGAAAGCATGTGTGAGAGACCCCAAGAGTCCCTGTGTG 599

OY 801 CTGGCAGTTGGGGGTGGGGGTCTCTGTGCGCGGGGTGTGTGCTGCTGTGAGGTGGG 860

Db 600 CTGGCAGTTGGGGGTGGGGGTCTCTGTGCGCGGGGTGTGTGCTGCTGTGAGGTGGG 659

OY 861 TGGGAGCATGTATCCCATATTTGTCATGAGAGCAATGGGCGACATGCTTCATATGGGGC 920

Db 660 TGGGAGCATGTATCCCATATTTGTCATGAGAGCAATGGGCGACATGCTTCATATGGGGC 716

OY 921 ATCAGAGCGGCGAAGCTGTCTACACTTCCAGACATCAGAGTGTGGCCAAAGAGCTGGGT 980

Db 717 ATCAGAGCGGCGAAGCTGTCTACACTTCCAGACATCAGAGTGTGGCCAAAGAGCTGGGT 775

OY 981 GCCAAGACGCTGGCGCTGCGGCGCTGTGAGTGCATGCAGG 1020

Db 776 G-CAGAGCGGTGGCGCTGCGGCGCTGTGAGTGCATGCAGG 814

RESULT 4

LOCUS BG749122 948 bp mRNA linear EST 15-MAY-2001

DEFINITION 602706734F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4843329.5,

ACCESSION BG749122

VERSION BG749122.1 GI:14059775

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS 1 (bases 1 to 948)

TITLE NIH-MGC http://mgi.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabps-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1CM1679 row: c column: 10
 High quality sequence stop: 802.

FEATURES

source
 1. 948
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4843329"
 /clone.lib="NIH_MGC_43"
 /tissue.type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pOTB7; Site: 1; XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GCCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

BASE COUNT 185 a 238 c 345 g 180 t
 ORIGIN

Query Match 49.6%; Score 736.4; DB 10; Length 948;
 Best Local Similarity 93.08; Pred. No. 1e-156;
 Matches 859; Conservative 0; Mismatches 56; Indels 9; Gaps 8;

OY 228 AGGAGCTGGGGCCACGCTGGTGGGATAGTTGGGAGGAGGCTGCTACCTGGTCTGC 287
 DB 2 AGGAGCTGGGGCCACGCTGGTGGGATAGTTGGGAGGAGGCTGCTACCTGGTCTGC 61
 OY 288 AGAATGAGAGCGCCCTGTGGGAGAGCATGCCAGACGAGGCCCTTTCACGTGTCAACCT 347
 DB 62 AGAATGAGAGCGCCCTGTGGGAGAGCATGCCAGACGAGGCCCTTTCACGTGTCAACCT 121
 OY 348 CTGTGGAGAGCTGGGGCTGTGCCAGGTGGCGGCAATGCTCTCTCTCAAGTGTGAG 407
 DB 122 CTGTGGAGAGCTGGGGCTGTGCCAGGTGGCGGCAATGCTCTCTCTCAAGTGTGAG 181
 OY 408 AATGTGAGAGCGCCCTGTGGGAGAGCATGCCAGATTCGGGCAATTCCTCCAGAGATG 467
 DB 182 AATGTGAGAGCGCCCTGTGGGAGAGCATGCCAGATTCGGGCAATTCCTCCAGAGATG 241
 OY 468 GCCAAGAGAGGATGACAGACCTGTGTCTCTCAAGGGGTAATCGGGGCAATCTGCT 527
 DB 242 GCCAAGAGAGGATGACAGACCTGTGTCTCTCAAGGGGTAATCGGGGCAATCTGCT 301
 OY 528 GCTATGCTGTAAGGAAGTGGGCAATTCCTGCACCATGCTCTCCGAGAGCACTCC 587
 DB 302 GCTATGCTGTAAGGAAGTGGGCAATTCCTGCACCATGCTCTCCGAGAGCACTCC 361
 OY 588 CTGCAAGTGTGTGAGAGGCTGTGAGGGGAGGGGCGGAGTTCACTGAGTGAAGGTC 647
 DB 362 CTGCAAGTGTGTGAGAGGCTGTGAGGGGAGGGGCGGAGTTCACTGAGTGAAGGTC 421
 OY 648 TGGGAGAGAGCCCAATCTGAGGGGCGCAAGAGTTGGCCAGAGGAGGAGGCTGGAGATTC 707
 DB 422 TGGGAGAGAGCCCAATCTGAGGGGCGCAAGAGTTGGCCAGAGGAGGAGGCTGGAGATTC 481
 OY 708 CCCCGTTTGAACACCCCTATATATGAAAGCCACCGCACCTGTGTGACAGAGCTGAAA 767
 DB 482 CCCCGTTTGAACACCCCTATATATGAAAGCCACCGCACCTGTGTGACAGAGCTGAAA 541
 OY 768 CGAGTGTGAGAGCCCAACAGAGTGTGTGTGTGAGTGGGGGGGCTGTCTCTG 827
 DB 542 CGAGTGTGAGAGCCCAACAGAGTGTGTGTGTGAGTGGGGGGGCTGTCTCTG 601

OY 828 GCCGGGTGTGTGC-TGGCTGTGTGAGAGTGGGTGGCAGCATGTACCATTCATTTGCCAT 886
 DB 602 AGCGGGGTGTGTGC-TGGCTGTGTGAGAGTGGGTGGCAGCATGTACCATTCATTTGC--A 659
 OY 887 GGAGACCATGGGACACCTGCTTCATTCGGGCGCATACACCGGCAAGTGTGTCACT 946
 DB 660 TGGAGACCATGGGACACCTGCTTCATTCGGGCGCATACACGCGGCAAGTGTGTCACT 717
 OY 947 TCCAGACATCACAGTGTGTGCGCCAGAGCTGTGTGCGCAAGAGCTGTGTGCGGCGCT 1006
 DB 718 TCCAGACATCACAGTGTGTGCGCCAGAGCTGTGTGCGGCGCATACAGTGTGTGCGGCGCT 775
 OY 1007 GGAATGATGAGTGTGTGCAAGATTCATCTGTGAAG-TGGTGGAGACACCGAGCTGTGA 1065
 DB 776 GAGTGTGATGAGTGTGTGCAAGATTCATCTGTGAAGTGGTGGAGACACCGAGCTGTGTG 835
 OY 1066 GCGCTGTGACAGAGCTCTGTGATGATGAGGCTGTGTGCGGCGCTGTGTGCGGCGAG 1125
 DB 836 AACGCTGTGACAGAGCTCTGTGATGATGAGGCTGTGTGCGGCGCATACAGATGACCTCTGTGTGAGCA 895
 OY 1126 CTTAGACCATCTCTCTGAGGCC 1149
 DB 896 CTTAG-TAGCATTTATATGAGGCC 918

RESULT 5
 LOCUS BG473546 807 bp mRNA linear EST 21-MAR-2001
 DEFINITION 602116083f1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4647572.5,
 mRNA sequence.
 ACCESSION BG473546
 VERSION BG473546.1 GI:13405821
 KEYWORDS EST,
 SOURCE human.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 807)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabps-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: L1CM1424 row: f column: 21
 High quality sequence stop: 802.

FEATURES

source
 1. 807
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4647572"
 /clone.lib="NIH_MGC_16"
 /tissue.type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pOTB7; Site: 1; XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GCCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

BASE COUNT 148 a 220 c 291 g 148 t

Query Match

48.9%; Score 726.4; DB 10; Length 807;


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image:3393393"
/clone_lib=Soares_NSf_F8_9w.orl.P.S1"
/lab_host=DH10B"
/node_Organ: pooled; Site: pT73D-Pac (Pharmacia) with
a modified polylinker; Site: 1: Not I; Site: 2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and six circles were made in vitro
following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NBHsp.pool.1:
309384-310919, 332308-3; Soares NBHsp.pool.1:
145033-147335, 147720-148103, 148872-149255, 15002-
150407, 151176-152323, Soares NBHSp_F8_9w.pool.1:
348280-760583, 772104-774407, Soares NBHsp.pool.1:
704776-306311, 350136-322823, 326280-32663 Soares NBHsp.
pool.1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fátima Bonaldo.
223 c 195 g 164 t

```

OY	777	AGGACCCACCAGAGTGCCTTGTCGTGCAGTGTGGGGTTGGGGGTCTCCTGCGGGGGT	836
Dd	732	AGTACCCCACAAAGSTGCTTAAGTACTGGCAATGTCCGTATAGAGGTCTCCTGCGGGGGT	836
OY	837	GTCGCTGGCCCTGCTGAGCTGTGGGCTGGAGAGATATTACCATTCCTCCATYGAGAGCCCAT	673
Dd	672	GTCGCTGGCCCTGCTGAGGTTGGGCTGAGAGATATTACCATTCATYTGGCATGAGAGCCAT	896
OY	897	GGGGCACACATGCTTAATTCGGCCATCAGACCGGCGAACCTGCTAATCTTCAGAGCCAT	613
Dd	612	GGGGCACACATGCTTAATTCGGCCATCAGACCTGGCAACCTGCTAATCTTCAGAGCCAT	956
OY	957	AACAGTGTGCCAAGAGCCTGGGTGGCCAAGAGGTGGCTGCTGGGGCCCTGGAGTAGAT	553
Dd	552	ACCAGTGTGCCAAGAGCCTGGGTGGCCAAGAGGTGGCTGCTGGGGCCCTGGAGTAGAT	1016
OY	1017	CAGGTGTCAAAGATTACTCTGGAAGTGTGGAGAACAACGAGGCTGTGAGCCGCTGGCAG	493
Dd	492	CAGGTGTCAAAGATTACTCTGGAAGTGTGGAGAACAACGAGGCTGTGAGCCGCTGGCAG	1076
OY	1077	CAGCTCCGATGATATAGCGTATGCTGTGTGAGAGCTGCTGCTGGGGCAGCCTTAACAGCC	433
Dd	432	CAGCTCCGATGATATAGCGTATGCTGTGTGAGAGCTGCTGCTGGGGCAGCCTTAACAGCC	1136
OY	1137	ATTCACACAGCCCTCCTGGAGAGGCTCCAGGCGAGAGGCTGCTGCCCCCTTCCCTGACT	373
Dd	372	ATTCACACAGCCCTCCTGGAGAGGCTCCAGGCGAGAGGCTGCTGCCCCCTTCCCTGACT	1196
OY	1197	TCAAGTGTGATAATGCTGTGTGAGAGCAACAACATCAACAGCCAGAGCTGAGAGCTTTG	313
Dd	312	TCAAGTGTGATAATGCTGTGTGAGAGCAACAACATCAACAGCCAGAGCTGAGAGCTTTG	1256
OY	1257	AAAACCCACATGGGCGAGCTGTAGAGGGTCCATCCTGGCCCAAAAGCCCTTGAAGGC	253
Dd	252	AAAACCCACATGGGCGAGCTGTAGAGGGTCCATCCTGGCCCAAAAGCCCTTGAAGGC	1316
OY	1317	CCATGGACAGTCTGTGTCTGATGAGAGAGCACTAGTGTGAGAGTGGCAGATGGAGAC	193
Dd	132	CCATGGACAGTCTGTGTCTGATGAGAGAGCACTAGTGTGAGAGTGGCAGATGGAGAC	1376
OY	1377	TGCCCTTGCAACGTGTGCTGGCTCCTCTAAGAGAGCCCTCTGAGCTGTCTTTTG	1436
Dd	132	TGCCCTTGCAACGTGTGCTGGCTCCTCTAAGAGAGCCCTCTGAGCTGTCTTTTG	1436
OY	1437	GCTTCGCAACATCGGGCAATAAAGACTTCTGAAATGAAAAAAA	1485
Dd	72	GCTTCGCAACATCGGGCAATAAAGACTTCTGAAATGAAAAAAA	24
RESULT	B		
LOCUS	BG774012		
DEFINITION	BG774012	667 bp mRNA linear EST 15-MAY-2001	
ACCESSION	602661747F1 NIH_MGC_21 Homo sapiens CDNA clone IMAGEF:4810186		
KEYWORDS	BG774012.1 GI:14044311		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	1 (bases 1 to 667)		
TITLE	NIH-MGC http://mgc.ncl.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
CONTACT:	Robert Stenmark		

found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L10M1665 row: n column: 11
 High quality sequence stop: 667.

FEATURES

Source
 1. 667
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4810186"
 /clone_lib="NIH.MGC.21"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: placenta; Vector: pOTB7; site_1: XhoI;
 site_2: EcoRI; cDNA made by oligo-dT priming.
 directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAGCAG(C). Size selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the Laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 116 a 176 c 251 g 124 t
 ORIGIN

Query Match 44.8%; Score 666; DB 10; Length 667;
 Best Local Similarity 100.0%; Pred. No. 8,7e-141;
 Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

209 CCCTGGCAGGCTGGAGGAGAGAGCTGGGGCCAGCTGCTGGGATGATTGGCAGGGA 268
 |||||||
 Db 2 CCCTGGCAGGCTGGAGGAGAGAGCTGGGGCCAGCTGCTGGGATGATTGGCAGGGA 61
 269 GCGTCTACCTGCTCCTCAGAAATGAGAGAGGCGCTGTGGCAGAGATGCCAGCAGAGCC 328
 |||||||
 Db 62 GCGTCTACCTGCTCCTCAGAAATGAGAGAGGCGCTGTGGCAGAGATGCCAGCAGAGCC 121
 329 CTTTACAGTGTGACACCTCTGTGTGAGAGAGTGGGCGCTGTCCAGTGGCGGATGCC 388
 |||||||
 Db 122 CTTTACAGTGTGACACCTCTGTGTGAGAGAGTGGGCGCTGTCCAGTGGCGGATGCC 181
 389 TGTCTCTCAAGTGTGAGAAATGTCACGCCGCGCTCTTCAGAGATTGGGGCATGGG 448
 |||||||
 Db 182 TGTCTCTCAAGTGTGAGAAATGTCACGCCGCGCTCTTCAGAGATTGGGGCATGGG 241
 449 GCATTTGTCAGAGAGATGGCCAGAGAGGATGACAGACCTGTGTGCTCTCAGGGGG 508
 |||||||
 Db 242 GCATTTGTCAGAGAGATGGCCAGAGAGGATGACAGACCTGTGTGCTCTCAGGGGG 301
 509 TAATGGGGGCATGCTGCTGCTATGCTGTAGAAAGCTGGGCATTCCTGCCACCATGCT 568
 |||||||
 Db 302 TAATGGGGGCATGCTGCTGCTATGCTGTAGAAAGCTGGGCATTCCTGCCACCATGCT 361
 569 GCTCCCGAGAGACCTCTCTGCAAGTGTGAGAGGCTGCAAGGGGGAGAGGGCCGAGGT 628
 |||||||
 Db 362 GCTCCCGAGAGACCTCTCTGCAAGTGTGAGAGGCTGCAAGGGGGAGAGGGCCGAGGT 421
 629 TCAGTGAAGTGAAGAGTCTGGAGAGAGCCCAATCTGAGGGGGCAAGAGTTGGCCAAAG 688
 |||||||
 Db 422 TCAGTGAAGTGAAGAGTCTGGAGAGAGCCCAATCTGAGGGGGCAAGAGTTGGCCAAAG 481
 482 GGACGGCTGGGAGATGTCCCCCTTTGACACCCCTTAATATGAGAAAGGCCAGCCAG 541
 689 GGACGGCTGGGAGATGTCCCCCTTTGACACCCCTTAATATGAGAAAGGCCAGCCAG 748
 |||||||
 Db 749 GGACGGCTGGGAGATGTCCCCCTTTGACACCCCTTAATATGAGAAAGGCCAGCCAG 808
 749 CCTGTGAGAGAGTGAAGAGTCTGAGAGATGCCACCATGCTGGCTGTGCTGTGGCAT 808
 |||||||
 Db 542 CCTGTGAGAGAGTGAAGAGTCTGAGAGATGCCACCATGCTGGCTGTGCTGTGGCAT 601
 809 TGGGGGTGGGGTCTCTGAGCGGGGGTGGTGGCTGGCTGGAGAGTGGCTGGCAGCA 868
 |||||||
 Db 602 TGGGGGTGGGGTCTCTGAGCGGGGGTGGTGGCTGGCTGGAGAGTGGCTGGCAGCA 661
 869 TGTACC 874

Db 662 TGTACC 667
 |||||||

RESULT 9

BE912245 716 bp mRNA linear EST 05-JUN-2001
 LOCUS 602806987F1 NCI.CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4939221
 DEFINITION 5', mRNA sequence.
 ACCESSION BE912245
 VERSION BE912245.1 GI:14292721
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 1 (bases 1 to 716)
 NIH-MGC http://mgc.nci.nih.gov/
 *AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 *TITLE Unpublished (1999)
 *JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs.rem@llnl.gov
 Tissue Procurement: David N. Louis, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L10M10875 row: n column: 22
 High quality sequence stop: 695.

FEATURES

Source

1. 716
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4939221"
 /clone_lib="NCI.CGAP_Brn67"
 /tissue_type="anaplastic oligodendroglioma with 1p/19q
 loss"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: brain; Vector: pCMV-SPORT6; site_1: NotI;
 site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2.3 kb. Constructed by Life
 Technologies. Note: this is a NCI.CGAP Library."
 BASE COUNT 137 a 208 c 230 g 141 t
 ORIGIN

Query Match 44.7%; Score 663.4; DB 10; Length 716;
 Best Local Similarity 99.4%; Pred. No. 3.5e-140;
 Matches 697; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

780 ACCCCACAGTGCCCTGTGCTGCGAGTGGGGGTCTCTGCGCGGGGTGGTG 839
 |||||||
 Db 1 ACCCCACAGTGCCCTGTGCTGCGAGTGGGGGTCTCTGCGCGGGGTGGTG 60
 840 GGTGGCTGCTGAGAGTGGCTGCGAGCATGTACCATATTGCCATGAGAGACCATGG 899
 |||||||
 Db 61 GGTGGCTGCTGAGAGTGGCTGCGAGCATGTACCATATTGCCATGAGAGACCATGG 120
 900 GCACACTGCTTAATGCGGCATTCACAGCGGCAAGCTGTGACACTTCACAGATCACC 959
 |||||||
 Db 121 GCACACTGCTTAATGCGGCATTCACAGCGGCAAGCTGTGACACTTCACAGATCACC 180
 960 ACTGTGGCCAAAGAGCTGGGTGCAAGAGCGGTGGCCGCTGGGGCCCTGGAGTGCATGCAG 1019
 |||||||
 Db 181 ACTGTGGCCAAAGAGCTGGGTGCAAGAGCGGTGGCCGCTGGGGCCCTGGAGTGCATGCAG 240
 1020 GGTGCAAGATTCACCTGTAAGTGGTGGAGACACCGAGAGCTGTGAGCGCTGTGACGCG 1079
 |||||||
 Db 241 GGTGCAAGATTCACCTGTAAGTGGTGGAGACACCGAGAGCTGTGAGCGCTGTGACGCG 300
 1080 CTCCTGATGATGAGCGATGCTGTGTGAGAGCTGCTGTGGGGCAGCCCTTACAGACCATC 1139

Query Match	Best Local Similarity	43.5%	Score 646.0	DB 9	Length 679
Query 809	TGGGGGTGGGGGTCCTCCCTGGCCGGGGGTGGCTGGCTGCTCTGAGGTGGGCTGGAGCA	668	TGGGGGTGGGGGTCCTCCCTGGCCGGGGGTGGCTGGCTGCTCTGAGGTGGGCTGGAGCA	609	
DB 668	TGGGGGTGGGGGTCCTCCCTGGCCGGGGGTGGCTGGCTGCTCTGAGGTGGGCTGGAGCA	609	TGGGGGTGGGGGTCCTCCCTGGCCGGGGGTGGCTGGCTGCTCTGAGGTGGGCTGGAGCA	609	
DB 869	TGTACCATCATATGTCATGAGAGAGCCATGGGGGACACACTGCTTCAATGCGGCATATCAGC	928	TGTACCATCATATGTCATGAGAGAGCCATGGGGGACACACTGCTTCAATGCGGCATATCAGC	928	
DB 608	TGTACCATCATATGTCATGAGAGAGCCATGGGGGACACACTGCTTCAATGCGGCATATCAGC	549	TGTACCATCATATGTCATGAGAGAGCCATGGGGGACACACTGCTTCAATGCGGCATATCAGC	549	
DB 929	CGGCAAGCTGTGTACATCTCCAGATATCAGACACTGCTGGGCGAAGAGCTGGGGGCGCAAGC	988	CGGCAAGCTGTGTACATCTCCAGATATCAGACACTGCTGGGCGAAGAGCTGGGGGCGCAAGC	988	
DB 548	CGGCAAGCTGTGTACATCTCCAGATATCAGACACTGCTGGGCGAAGAGCTGGGGGCGCAAGC	489	CGGCAAGCTGTGTACATCTCCAGATATCAGACACTGCTGGGCGAAGAGCTGGGGGCGCAAGC	489	
DB 989	GGTGGCCCTCTGGGGCCCTGGAGTGTATGCAAGTGTGCAAGATTTCACTCTGAAGTGTGA	1048	GGTGGCCCTCTGGGGCCCTGGAGTGTATGCAAGTGTGCAAGATTTCACTCTGAAGTGTGA	1048	
DB 488	GGTGGCCCTCTGGGGCCCTGGAGTGTATGCAAGTGTGCAAGATTTCACTCTGAAGTGTGA	429	GGTGGCCCTCTGGGGCCCTGGAGTGTATGCAAGTGTGCAAGATTTCACTCTGAAGTGTGA	429	
DB 1049	GGACACCGAGGGCTGTAGAGGCTGTGAGAGCTCTCTGGATATGATGAGCTATCTGTGTGA	1108	GGACACCGAGGGCTGTAGAGGCTGTGAGAGCTCTCTGGATATGATGAGCTATCTGTGTGA	1108	
DB 428	GGACACCGAGGGCTGTAGAGGCTGTGAGAGCTCTCTGGATATGATGAGCTATCTGTGTGA	369	GGACACCGAGGGCTGTAGAGGCTGTGAGAGCTCTCTGGATATGATGAGCTATCTGTGTGA	369	
DB 1109	GGCTGCTCTGGGGGAGAGCTTATGACAGCCATATCTAGAGCTCTCTGGAGGCTCTCAGGC	1168	GGCTGCTCTGGGGGAGAGCTTATGACAGCCATATCTAGAGCTCTCTGGAGGCTCTCAGGC	1168	
DB 368	GGCTGCTCTGGGGGAGAGCTTATGACAGCCATATCTAGAGCTCTCTGGAGGCTCTCAGGC	309	GGCTGCTCTGGGGGAGAGCTTATGACAGCCATATCTAGAGCTCTCTGGAGGCTCTCAGGC	309	
DB 1169	CGAGGGCTCTGCTGCCCCCTTCCCTGACTTCACTTGTGTGTAATCTGTGTGTGAGGACAA	1228	CGAGGGCTCTGCTGCCCCCTTCCCTGACTTCACTTGTGTGTAATCTGTGTGTGAGGACAA	1228	
DB 308	CGAGGGCTCTGCTGCCCCCTTCCCTGACTTCACTTGTGTGTAATCTGTGTGTGAGGACAA	249	CGAGGGCTCTGCTGCCCCCTTCCCTGACTTCACTTGTGTGTAATCTGTGTGTGAGGACAA	249	
DB 1229	CATCAACACCGAGAGAGCTGAGGCTTTGAAACCCACCTGGGCGAGAGCTGTGAGGGGTCC	1288	CATCAACACCGAGAGAGCTGAGGCTTTGAAACCCACCTGGGCGAGAGCTGTGAGGGGTCC	1288	
DB 248	CATCAACACCGAGAGAGCTGAGGCTTTGAAACCCACCTGGGCGAGAGCTGTGAGGGGTCC	189	CATCAACACCGAGAGAGCTGAGGCTTTGAAACCCACCTGGGCGAGAGCTGTGAGGGGTCC	189	
DB 1289	ATCTGTGGCCCCAAAGACCCCTGAGAGGCCATATGACAGCTCTGTGTGTGATGAGAGAGA	1348	ATCTGTGGCCCCAAAGACCCCTGAGAGGCCATATGACAGCTCTGTGTGTGATGAGAGAGA	1348	
DB 188	ATCTGTGGCCCCAAAGACCCCTGAGAGGCCATATGACAGCTCTGTGTGTGATGAGAGAGA	129	ATCTGTGGCCCCAAAGACCCCTGAGAGGCCATATGACAGCTCTGTGTGTGATGAGAGAGA	129	
DB 1349	CTCAGTGTCTGGAGATGAGAGAGAGCTGAGGCTTGTGACACTGTCTGGGCTGCTCTCGA	1408	CTCAGTGTCTGGAGATGAGAGAGAGCTGAGGCTTGTGACACTGTCTGGGCTGCTCTCGA	1408	
DB 128	CTCAGTGTCTGGAGATGAGAGAGAGCTGAGGCTTGTGACACTGTCTGGGCTGCTCTCGA	69	CTCAGTGTCTGGAGATGAGAGAGAGCTGAGGCTTGTGACACTGTCTGGGCTGCTCTCGA	69	
DB 1409	AGGAAGCCTCTCTGAGACTGTCTTTTGTGGCTCTCGAGAACTCGGCGCAATTAACACTTT	1468	AGGAAGCCTCTCTGAGACTGTCTTTTGTGGCTCTCGAGAACTCGGCGCAATTAACACTTT	1468	
DB 68	AGGAAGCCTCTCTGAGACTGTCTTTTGTGGCTCTCGAGAACTCGGCGCAATTAACACTTT	9	AGGAAGCCTCTCTGAGACTGTCTTTTGTGGCTCTCGAGAACTCGGCGCAATTAACACTTT	9	
DB 1469	CTGAATTG.1476		CTGAATTG.1476		
DB 8	CTGAATTG.1		CTGAATTG.1		
RESULT 11					
LOCUS	AI741818	666 bp	mRNA	linear	EST 19-DEC-1999
DEFINITION	w929c09.x1 Scores_NSF_F8_9M_OT_PA.P_S1 Homo sapiens CDNA clone IMAGE:2366512 3' similar to SW:SDHL_RAT P09367 L-SERINE DEHYDRATASE				
ACCESSION	AI741818				
VERSION	AI741818.1				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryotic Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.				
AUTHORS	NCI-CCAP http://www.ncbi.nlm.nih.gov/ccap.				
TITLE	National Cancer Institute, Cancer Genome/Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Robert Strausberg, Ph.D.				
	Email: cgapds1@nsl.nih.gov				


```

Db 545 CAAGCTGTACACTTCCACATACACACAGCTGTGGCCAAAGAGCTGGGTCCCAAGACGCT 486
OY 992 GGGCGCTGGGCGCTGTGAGTGCATGAGTGTGCAAGATTCCTCTAAGTGTGGAGGA 1051
Db 485 GGGCGCTGGGCGCTGTGAGTGCATGAGTGTGCAAGATTCCTCTAAGTGTGGAGGA 426
OY 1052 CACGAGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 1111
Db 425 CACGAGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 366
OY 1112 TGCCTGTGGGAGACCTTTAGACACCATCTACTAGGCTCTGAGGCTCTGAGGCTCTGAGGCT 1171
Db 365 TGCCTGTGGGAGACCTTTAGACACCATCTACTAGGCTCTGAGGCTCTGAGGCTCTGAGGCT 306
OY 1172 GGGCTGTGGGCGCTGTGAGTGCATGAGTGTGCAAGATTCCTCTAAGTGTGGAGGA 1231
Db 305 GGGCTGTGGGCGCTGTGAGTGCATGAGTGTGCAAGATTCCTCTAAGTGTGGAGGA 246
OY 1232 CAACAGCCGAGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 1291
Db 245 CAACAGCCGAGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 186
OY 1292 CTGGCGCCCAAGACCCCTGAGAGGCGCCATGAGAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGT 1351
Db 185 CTGGCGCCCAAGACCCCTGAGAGGCGCCATGAGAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGT 126
OY 1352 AGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1411
Db 125 AGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 66
OY 1412 AAGCCCTCTGAGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 1471
Db 65 AAGCCCTCTGAGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 6
OY 1472 AATGG 1476
Db 5 AATGG 1

```

RESULT 13
BE273896 645 bp mRNA linear EST 13-JUL-2000
LOCUS BE273896 601104523f1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:347525 5'
DEFINITION mRNA sequence.
ACCESSION BE273896
VERSION BE273896.1 GI:9148832
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS 1 (bases 1 to 645)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs.fda.gov
Tissue Procurement: DCTD/DP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L10M138 row: n column: 06
High quality sequence stop: 626.
Location/Qualifiers
1. 645
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:347525"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"

```

BASE COUNT 113 a 168 c 239 g 125 t
ORIGIN
Query Match 40.6%; Score 603; DB 10; Length 645;
Best Local Similarity 98.9%; Pred. No. 1.7e-126;
Matches 628; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
OY 225 GGCAGAGAGTGGGGCCACGCTGTGCTGGATAGTGGGCGAGG-AGGCTGTACCTGCT 283
Db 1 GGCAGAGAGTGGGGCCACGCTGTGCTGGATAGTGGGCGAGTGTAGCTGTCTACCTGCT 60
OY 284 CTCCAGATGAGAGCGCCCTGTGAGAGCATGCGCAAGACAGAGACCCCTTACAGTGTGAC 343
Db 61 CTCCAGATGAGAGCGCCCTGTGAGAGCATGCGCAAGACAGAGACCCCTTACAGTGTGAC 120
OY 344 ACCTCTGTGAGAGAGCTGGGCGCTGTGCCAGTGGCGGCGATGCTGTCTCTCAAGTG 403
Db 121 ACCTCTGTGAGAGAGCTGGGCGCTGTGCCAGTGGCGGCGATGCTGTCTCTCAAGTG 180
OY 404 TGAGATGTGACACCCAGCGGCTCTTCAAGATTCGGGCGATTCGATTCGCGAGGA 463
Db 181 TGAGATGTGACACCCAGCGGCTCTTCAAGATTCGGGCGATTCGATTCGCGAGGA 240
OY 464 GATGCGCAAGAGGATGACAGACCTGTGTCTCTCAAGGGGTATGCGGCGATTCG 523
Db 241 GATGCGCAAGAGGATGACAGACCTGTGTCTCTCAAGGGGTATGCGGCGATTCG 300
OY 524 TGCTGCTATGCTGTGAGAGAGCTGGGCAATCTCTGACACATGCTGCTCCGAGAGCA 583
Db 301 TGCTGCTATGCTGTGAGAGAGCTGGGCAATCTCTGACACATGCTGCTCCGAGAGCA 360
OY 584 CTCCCTGACAGTGTGACAGAGCTGTGAGAGGAGGAGGCGGAGAGTTCACCTGATGAGAA 643
Db 361 CTCCCTGACAGTGTGACAGAGCTGTGAGAGGAGGAGGCGGAGAGTTCACCTGATGAGAA 420
OY 644 GGTCTGGAGAGGCGCAATCTGAGGCGCAAGAGTGGCCAGAGGAGCGCTGGAGAA 703
Db 421 GGTCTGGAGAGGCGCAATCTGAGGCGCAAGAGTGGCCAGAGGAGCGCTGGAGAA 480
OY 704 TGTCCCGCGTTTGACACCCCTTAATATGAAAGGCCAGCGAGCTGTGAGAGAGCT 763
Db 481 TGTCCCGCGTTTGACACCCCTTAATATGAAAGGCCAGCGAGCTGTGAGAGAGCT 540
OY 764 GAAAGCAGTGTGAGAGACCCCAACAGTGTGCTGTGAGAGTGTGGGGGTGGGCTGT 823
Db 541 GAAAGCAGTGTGAGAGACCCCAACAGTGTGCTGTGAGAGTGTGGGGGTGGGCTGT 600
OY 824 CCTGGCGGGGGGTGGGCTGTGAGAGTGTGGGCTGTGAGAGTGTGGGCTGTGAGAGTGTG 858
Db 601 CCTGTGCGGGGGGTGGGCTGTGAGAGTGTGGGCTGTGAGAGTGTGGGCTGTGAGAGTGTG 634

```

RESULT 14
BG169688 964 bp mRNA linear EST 06-FEB-2001
LOCUS BG169688 602324674f1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4427490 5'
DEFINITION mRNA sequence.
ACCESSION BG169688
VERSION BG169688.1 GI:12676391
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 964)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps@email.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LLM10176 row: d column: 19
 High quality sequence stop: 716.

FEATURES
 source location/Qualifiers
 1..964

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4427490"
 /clone_lid="NIH MGC 89"
 /tissue_type="hypernephroma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: kidney; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."
 BASE COUNT 210 a 233 c 360 g 161 t
 ORIGIN

Query Match 39.5%; Score 587.2; DB 10; Length 964;
 Best Local Similarity 95.6%; Pred. No. 7.3e-123;
 Matches 658; Conservative 0; Mismatches 23; Indels 7; Gaps 5;

221 GGAGGCGAGGAGCTGGGCGCAGCGCTGCTGGAGATTGGGAGAGGCTGTACT 280
 1 GGAGGCGAGGAGCTGGGCGCAGCGCTGCTGGAGATTGGGAGAGGCTGTACT 59
 281 GGTCTCCAGATTGACGCGCTGTGGCAGAGCAGTCCAGAGAGAGCCCTTACAGTGT 340
 60 GGTCTCCAGATTGACGCGCTGTGGCAGAGCAGTCCAGAGAGAGCCCTTACAGTGT 119
 341 CACACCTCTGTTGAGACCTGGGCGCTGCCAGTGGCGGCGATGCTGTCTCTCAA 400
 120 CACACCTCTGTTGAGACCTGGGCGCTGCCAGTGGCGGCGATGCTGTCTCTCAA 179
 401 GTGTGAGATTGACGCGCTGCCAGTGGCGGCGATGCTGTCTCTCAA 460
 180 GTGTGAGATTGACGCGCTGCCAGTGGCGGCGATGCTGTCTCTCAA 239
 461 GGAGATGGCCAAAGAGGATGACACCTGCTGCTCTCAAGGGGGTAAATGCGGGCAT 520
 240 GGAGATGGCCAAAGAGGATGACACCTGCTGCTCTCAAGGGGGTAAATGCGGGCAT 299
 521 CGCTGCTCTTATGCTGTAGGAAGCTGGGCAATCTCTGACATGCTGCTCCCGAAG 580
 300 CGCTGCTCTTATGCTGTAGGAAGCTGGGCAATCTCTGACATGCTGCTCCCGAAG 359
 581 CACCTCCTGAGTGTGACAGAGGCTGACAGGGGAGGGGCGGAGTTGACATGCTG 640
 360 CACCTCCTGAGTGTGACAGAGGCTGACAGGGGAGGGGCGGAGTTGACATGCTG 419
 641 AAAGTCTGGAAGAGGCAATCTGAGGGCGCAAGAGTTGGCCAGAGGAGCGCTGGA 700
 420 AAAGTCTGGAAGAGGCAATCTGAGGGCGCAAGAGTTGGCCAGAGGAGCGCTGGA 479
 701 GAATGTCCTCCCGTGTGACACCCCTTAATATGAAAGGCCAGCGCAGCTGTGTGAGA 760
 480 GAATGTCCTCCCGTGTGACACCCCTTAATATGAAAGGCCAGCGCAGCTGTGTGAGA 539

QY 761 GCTGAAGCAGTGTGAGAGACCCACAGAGTCCCTGTGCTGCGAGTGGGGTGGGG 820
 DB 540 GCTGAAGCAGTGTGAGAGACCCACAGAGTCCCTGTGCTGCGAGTGGGGTGGGG 598
 QY 821 TCTCTGCGCGGGGTGCTGCTGCTGAGAGTGGGCTGGCAGCATGTACCATAT 880
 DB 599 TCTCTGCGCGGGGTGCTGCTGCTGAGAGTGGGCTGGCAGCATGTACCATAT 653
 QY 881 TCCCATGAGAGACCCATGAGGCGGAGGCGGAGTGG 908
 DB 654 TCCCATGAGAGACCCATGAGGCGGAGGCGGAGTGG 681

RESULT 15
 AA573827/c 614 bp mRNA linear EST 12-SEP-1997
 LOCUS Similar to SW:SDHL_RAT P09367 L-SERINE DEHYDRATASE ; mRNA
 DEFINITION

ACCESSION AA573827
 VERSION AA573827
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 614)
 AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/nciccap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps@email.nih.gov
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Stratagene, Inc., David B. Krizman, Ph.D.

CDNA Library Arraying: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/bbrp/image/image.html
 Insert Length: 1319 Std Error: 0.00
 Seq primer: -40m13 fwd. 5' from Amersham
 High quality sequence stop: 445.
 Location/Qualifiers
 1..614

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1012877"
 /clone_lid="NCI-CCAP_C02"
 /tissue_type="tumor"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: colon; Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Bulk colon villous adenoma. 5' adaptor sequence: 5' GAATTCGGGACAGAG 3' 3' adaptor sequence: 5' CTGCACTTTTCTTTTCTTTT 3' Average insert size: 1.1 kb."
 BASE COUNT 126 a 185 c 179 g 124 t
 ORIGIN

Query Match 39.3%; Score 583.6; DB 9; Length 614;
 Best Local Similarity 98.7%; Pred. No. 4.2e-122;
 Matches 610; Conservative 0; Mismatches 4; Indels 4; Gaps 2;

QY 861 TGGCAGCATGTATCCCATCTTTCATGTGAGAGCCATGGGGCAGCTGCTTCAATGCGGCC 920
 DB 614 TGGCAGCATGTATCCCATCTTTCATGTGAGAGCCATGGGGCAGCTGCTTCAATGCGGCC 555
 QY 921 ATCACAGCGGCGAGCTGTACACTTCCAGACATACCACTGTGTGGCCAGAGCGCTGGGT 980
 DB 554 ATCACAGCGGCGAGCTGTACACTTCCAGACATACCACTGTGTGGCCAGAGCGCTGGGT 495

Thu Jun 6 11:33:59 2002

us-09-088-435-2.rst

Page 12

[illegible]

Search completed: June 5, 2002, 00:42:48
Job time: 1836 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2002, 00:15:38 ; Search time 215.47 Seconds

(without alignments)
11832.821 Million cell updates/sec

Title: US-09-088-435-2

Perfect score: 1485

Sequence: 1 GGGTCGACCAACCGTCGATG.....TTTCTGATTTGAAAAAAA 1485

Scoring table: IDENTITY_NIC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match: 0%

Maximum Match: 100%

Listing first 45 summaries

Database :
1: N_Geneseq_032802.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1485	100.0	1485	22	AA512098	Human serine dehyd
2	1320	88.9	1364	22	AA541108	CDNA encoding nove
3	1315	88.6	1396	22	AA541634	CDNA encoding nove
4	1303.4	87.8	1379	22	AA157992	Human polynucleoti
5	1284.6	86.5	1304	22	AA159778	Human polynucleoti
6	1283	86.4	1327	24	ABA02877	Human DHX encoding
7	986.8	66.5	990	24	ABA02878	Human DHX coding s
8	374	25.2	390	20	AA541006	Human secreted pro
9	211.4	14.2	213	15	AA057448	Ser dehydrogenase

10	102	6.9	443	22	AA192045	Human polynucleoti
11	69.2	4.7	1545	19	AAV05037	Escherichia coli t
12	69.2	4.7	1545	19	AAV05038	Escherichia coli t
13	69.2	4.7	1545	19	AAV05039	Escherichia coli t
14	69.2	4.7	1545	20	AA224936	Wild type E.coli t
15	69.2	4.7	1545	20	AA224939	Mutant E.coli t
16	69.2	4.7	1545	20	AA224940	Mutant E.coli t
17	69.2	4.7	1545	20	AA224941	Mutant E.coli t
18	69.2	4.7	1545	20	AA208783	Mutagenised Escher
19	69.2	4.7	1545	20	AA208785	Wild-type Escheric
20	69.2	4.7	1545	20	AA208788	Mutagenised Escher
21	69.2	4.7	1545	20	AA208789	Mutagenised Escher
22	69.2	4.7	1545	21	AA208784	E. coli tlvA219 ge
23	69.2	4.7	1545	21	AA208788	E. coli tlvA219 ge
24	69.2	4.7	1545	21	AA208789	E. coli tlvA219 ge
25	69.2	4.7	1545	21	AA208784	E. coli tlvA219 ge
26	69.2	4.7	1545	22	AA208785	Escherichia coli t
27	69.2	4.7	1545	22	AA208788	Escherichia coli t
28	69.2	4.7	1545	22	AA208789	Escherichia coli t
29	69.2	4.7	1545	22	AA208783	Escherichia coli t
30	65.4	4.4	1189	23	ABL18395	Drosophila melanog
31	65.4	4.4	1447	23	ABL18395	Drosophila melanog
32	59.2	4.0	4403765	22	AA199683	Mycobacterium tube
33	59.2	4.0	4411529	22	AA199682	N. meningitidis pa
34	56.6	3.8	9243	21	AA541534	Genomic fragment
35	56.6	3.8	349980	21	AA541534	Genomic fragment
36	56.6	3.8	55909	22	AA541534	Genomic fragment
37	54.6	3.7	534720	19	AAV30458	Rhizobium species
38	54.6	3.7	536165	19	AAV30459	Rhizobium species
39	53.6	3.6	1509	20	AA211199	Mutant threonine d
40	53.6	3.6	1509	20	AA211199	Mutant threonine d
41	53.6	3.6	1509	20	AA211199	Mutant threonine d
42	53.6	3.6	1509	20	AA211199	Mutant threonine d
43	53.6	3.6	1620	20	AA211200	Mutant threonine d
44	53.6	3.6	1620	20	AA211200	Mutant threonine d
45	53.6	3.6	1638	20	AA211205	Mutant threonine d

ALIGNMENTS

RESULT 1
AA512098
ID AA512098 standard; DNA: 1485 BP.

XX AA512098;

DT 21-NOV-2001 (first entry)

DE Human serine dehydratase homologue (SDHH) DNA.

XX Serine dehydratase homologue; SDHH; human; cytosolic; antidiabetic; ds;
XX antithyroid; antilipemic; anorectic; nootropic; dermatological; cancer;
XX hepatocytic; metabolic; Addison's disease; cystic fibrosis; diabetes;
XX fatty hepatocirrhosis; galactosaemia; hyperadrenism; hypoadrenism;
XX hypoparathyroidism; hyperparathyroidism; hypercholesterolemia; goitre;
XX hyperthyroidism; hypothyroidism; hyperlipidaemia; hyperlipaemia; breast;
XX lipid myopathy; obesity; lipodystrophy; phenylketonuria; renal failure;
XX adenocarcinoma; leukaemia; lymphoma; melanoma; adrenal gland; bladder;
XX bone marrow; brain; cervix; heart; kidney; liver; muscle; ovary;
XX penis; prostate; pancreas; skin; spleen; testis; thymus; thyroid; uterus.

OS Homo sapiens.

XX key Location/Qualifiers

XX CDS 291..1280
XX /*tag= a
XX /product= "Human serine dehydratase homologue"
XX /transl_except= (pos:885..896, aa:Met-His)
XX /transl_except= (pos:1125..1136, aa:Ala-Ala)

XX US6277619-B1.

CC The sequence represents DNA encoding a human serine dehydratase homologue
CC (SDH) polypeptide. SDH, its fragments or its derivatives are useful for
CC treating or preventing a disorder of metabolism such as Addison's
CC disease, cystic fibrosis, diabetes, fatty hepatocarcinosis, galactosaemia,
CC goitre, hyper and hypoadrenalism, hyper and hypoparathyroidism,
CC hypercholesterolaemia, hyper and hypothyroidism, hyperlipidaemia,
CC hyperlipoproteinaemia, lipid myopathies, obesity, lipodystrophies,
CC phenylketonuria and renal failure, cancer, including adenocarcinoma,
CC leukaemia, lymphoma, melanoma and cancers of the adrenal gland, bladder,
CC bone, bone marrow, brain, breast, cervix, kidney, liver, muscle,
CC ovary, penis, prostate, pancreas, skin, spleen, testis, thymus, thyroid
CC and uterus.
XX
50 Sequence 1485 BP, 284 A, 405 C, 504 G, 292 T, 0 other.

[illegible][illegible]

DT 17-DEC-2001 (first entry)
XX cDNA encoding novel human enzyme polypeptide #324.
XX
XX Human: oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW lysase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotoxic; anticoagulant; ss.
XX
OS Homo sapiens.
XX
XX MO200155301-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01239.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0234999.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241788.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246533.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.

DT 17-DEC-2001 (first entry)
XX
DE cDNA encoding novel human enzyme polypeptide #850.
XX
KW Human: oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW lyase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotoxic; anticoagulant; ss.
XX
OS Homo sapiens.
XX
PN W020015301-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01239.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180638.
PR 02-MAR-2000; 2000US-0184664.
PR 16-MAR-2000; 2000US-0186350.
PR 17-MAR-2000; 2000US-0189874.
PR 18-APR-2000; 2000US-0190076.
PR 19-MAY-2000; 2000US-0198123.
PR 07-JUN-2000; 2000US-0205515.
PR 28-JUN-2000; 2000US-0209467.
PR 30-JUN-2000; 2000US-0214886.
PR 07-JUL-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217486.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234937.
PR 25-SEP-2000; 2000US-0234999.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246522.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.

PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251836.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR MPI: 2001-465566/50.
XX
DR P-PSDB; AAU23764.
XX
PT Novel polypeptides and polynucleotides useful for diagnosing,
PT preventing, treating neural, immune system, muscular, reproductive,
PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous
PT diseases
XX
PS Claim 4; SEQ ID NO 860; 1180bp; English.

polypeptides (e.g. A022315-A02314), and the cDNA and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention may comprise the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS), autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. influenza). The A0450785-A045184 represent cDNA sequences encoding for the novel human enzyme polypeptides of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WFO at [ftp.wipo.int/pub/publicated_pcl_sequences](http://wipo.int/pub/publicated_pcl_sequences).

sequence 1396 BP; 2/6 A; 380 C; 471 G; 267 T; 2 other,

Query Match	88.6%;	Score 1315;	DB 22;	Length 1366;
Best Local Similarity	99.8%;	Pred No. 5.6e-310;		
Matches 1326;	Conservative	1;	Mismatches	1;

OY	158	GCCGGAAAGCTGACCTCG- AAGTGTGACAGAGCCGGTCTTCGTGAAAGACCTGATTCCTTGACA	211
Db	39	gcccgaagctcgtcctgnaagctctgnaacagagccggtcttcctgnaaagagctgttccctggca	98
OY	217	GCGTGAAGGCGACGAGACTGGGGCCACGCTGGTCTGGAGTATGTTGGGCGACGAGAGCGTGTCT	276
Db	99	ggtctgagagcgcaagagctctgggcccagctgtgtctcggatatagtcttgagcagagagagctcgtct	155
OY	277	ACCTGGTCTCCAAATAGAGAGCGCCCTGTGGCAGAGCATGTCGCAAGCGAGAGCCCTTTCACG	336
Db	159	acctggtctcccaagaaatgagacgagccctctgtgaaagatcatcgacaagcagagagcccttcaag	218
OY	337	TGGTCACACCTCTGTTTGGAGAGCTGGGGCGCGTCCGAGGAGCGGCGGATGCGCTGTCTTC	396
Db	219	tgtgtcaacctctgtcttgtagagagctgaggcgctgtcccaagctggcgagagatgtccttcttc	278
OY	397	TCAAGTGTGAGAAATGTGCAGCCCGACGGCGCTTCCTCAAGATTGGGGCGATTGGCATTTCT	456
Db	279	tcaagctgtgagaaatgttgtagaccagcgagcccttcaagattctggggcgatttggagatttct	338
OY	457	GCCAGAGATGGCCAAAGAGGATGCAGACACTGGTGTGCTCCAGGGGATATTCGG	516

[illegible]

ID AI157992 standard; cDNA: 1379 BP.
XX
AC AA157992:
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 195.
XX
KW Human; nocitropic; immunosuppressant; cyostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokine; thrombolytic; drug screening; arthritis; inflammation;
leukaemia; ss.
XX
OS Homo sapiens.
XX
PN MO20015312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PI (HYSE-) HYSEQ INC.
PI
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR P-PSDB: AAM38836.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 195; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AI157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM4213) with nocitropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokine activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
CC assays for receptor activity, arthritis and inflammation, Leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 1379 BP; 264 A; 382 C; 468 G; 265 T; 0 other;

Query Match 87.8%; Score 1303.4; DB 22; Length 1379;
Best Local Similarity 99.9%; Pred. No. 3,76-307;
Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 181 CAGAGCCGCTTCGGAAGAGCTGCTCCCTGCAGAGCTGGAGGCGAGAGCTGGGCGC 240
72 cagagccgcttcctcgaagagctgctccctgcagagctggagggcagagagctgggccc 131

241 ACGCTGCTCGGATGATGTTGGGAGGAGGCTGTCTACCTGCTTCAGAAATGAGCGGC 300
132 aagcgtcctcggatgattggtggaagagagctgctacacgtgctccagaaagagcgc 191
301 CTGTGGCAGAGCATGCCAAGCAGAGAGCCCTTTCACGTGGTACACCTCTGTTGAGACT 360
192 ctgtggcagagcatgcccaagcagagagcccttcacgtgtgtaacacactgtttggagact 251
361 GGGCGCTTCCAGAGTGGCGGGGATGCTGTCTTCTTCAAGTGTGAGAAATGTGACGCCA 420
252 gggcgcttccagagtgccgagcgtgctgtcttcccaagtgtagaagtgagcgcga 311
421 GCGGCTCTTCAGATTCGGGGCATTTGGGCAATTTGTCAGAGATGAGCAAGAGGAT 480
312 gcggtccttcaagatctcgaggatctggatcttcgccaagagatgccaagagagat 371
481 GCAGACACCTGCTGCTCTCAGAGGCTAATGCGGCGATGCTGCTCTATGCTGCTA 540
372 gcagacacctgtgtgtcctcccaagggtaatgsggcatgctgtgctgtgctgcta 431
541 GGAAGCTGGCATTTCTGCTCAGACATGCTGCTCCCGAGAGACCTCCCTGCAAGTGTGC 600
432 ggaagctggcatcttctccacacatcgtctcccgagagacacctccctgaagtgctgc 491
601 AGAGCTGTCAGAGGAGGAGGCGCGAGTTTCAGCTGACTGGAAGGTTGGGACGAGGCA 660
492 agagctgcaagggagggagggagggagggagggagggagggagggagggagggaggg 551
661 ATCTGAGGCGCCAGAGCTGGGCGCAAGAGGAGGAGGCTGGGAGAAATGCTCCCGCTTGAAC 720
552 atctgagggcgcaagagatgttgccaagagggagggagggagggagggagggaggg 611
721 ACCCTCTAATATGAAAGGCCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
612 accctctaataatgaaaggccacgctgctgctgctgctgctgctgctgctgctgctg 671
781 CCCGACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
672 cccgacagagggctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 731
841 CTGGCTGCTGAGAGTGGGCTGGGAGAGTATGACCATATTCATGAGAGAGCCATGGGG 900
732 ctggctgctgagagtgggctgggagagcatgtacatcatgtgcatgagagagagag 791
901 CACACTGCTTAATGCGGCTCATACAGCGGAGAGCTGCTACACTTCACAGATCACCA 960
792 cacactgcttaatgctgctcatcacagcagcagcagcagcagcagcagcagcagcagc 851
961 GTGTGCGCAGAGGCTGGGTGCGAAGAGGAGGCGGCTCGGGCCCTGGAGAGCATGAGG 1020
852 gtgtgccaagagcctgggtgccaagagcagcagcagcagcagcagcagcagcagcagc 911
1021 TGTGCAAGATTCACCTCAAGTGTGAGAGACACCGAGGCTGTGAGCCCTGTGACGAGC 1080
912 tgtgcaagaattcacctcaagtgtgagagacacacagcagcagcagcagcagcagcagc 971
1081 TCTGTGATGATGAGGCTATGCTGTGAGGCTGCTGTGAGGAGCTTACAGAGCATCT 1140
972 tctgtgatgattgagcctatgctgtgagagcctgtgtgagagcagcagcagcagcagc 1031
1141 ACTGAGGCTGCTGCGAGGCTCCAGGCGGAGGCTGCTGCGGCTGCTGCTGCTGCTGCT 1200
1032 actgagcctcctcggagagctcagagcagcagcagcagcagcagcagcagcagcagc 1091
1201 TTGTGTATCATGCTGTGTGAGAGCAACAACATCAACAGCGAGAGCTGAGAGCTTTGA 1260
1092 ttgtgtatcatgctgtgtgagagcaacaacatcaacacagcagcagcagcagcagcagc 1151
1261 CCCACTGAGGCGAGTGTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
1152 cccactgagcaggtgtgagagggctcctcctcctcctcctcctcctcctcctcctcctc 1211

	indels	Gaps
195 GGAAGACGCTCCCTTCCCGCAGCGGGG		

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

OY 257 GTTGGCAGGAGGCTGCTACCTGATC-TCCAGATGAGCGGCTGAGGAGACCATG 315
 DB 64 GTTGGCAGGAGGCTGCTACCTGATC-TCCAGATGAGCGGCTGAGGAGACCATG 123
 OY 316 CCAAGCAGGAGGCTGCTACCTGATC-TCCAGATGAGCGGCTGAGGAGACCATG 375
 DB 124 CCAAGCAGGAGGCTGCTACCTGATC-TCCAGATGAGCGGCTGAGGAGACCATG 183
 OY 376 TGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 435
 DB 184 TGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
 OY 436 TGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 495
 DB 244 TGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 303
 OY 496 GGTCTCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 555
 DB 304 GGTCTCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363
 OY 556 CTTGGCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 582
 DB 364 CTTGGCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 390

RESULT 9

AA057448 ID AA057448 standard: cDNA to mRNA: 213 BP.

AC AA057448;
 DT 19-OCT-1994 (first entry)

DE Ser dehydrogenase like protein.

KM Human cDNA: library: enzyme: protein; ss.

OS Homo sapiens.

PN WO9403599-A1

PD 17-FEB-1994.

PF 04-AUG-1993; 93WO-JP01095.

PR 04-AUG-1992; 92JP-0208077.

PR 13-NOV-1992; 93JP-0327619.

PR 26-FEB-1993; 93JP-0061431.

PA (SAGA) SAGAMI CHEM RES CENTRE.

PI Iwahori A, Kato S, Kato T, Kim N, Oh S, Isekine S.

DR WPI: 1994-065688/08.

DR P-PSDB: AAR46109.

PT cDNA of human origin and proteins coded by it - which may be

PT expressed by in vivo or in vitro translation using sense RNA or

PT antisense DNA corresponding to the cDNA.

PS Claim 1; Page 68; 167pp; Japanese.

XX mRNA expressed in human fibrosarcoma cell line HUV-78 was

CC isolated and used to construct a cDNA library using vector

CC PT218R1. Clone HP00157 encoding Ser dehydrogenase-like protein

CC was isolated.

XX Sequence 213 BP; 34 A; 62 C; 77 G; 40 T; 0 other;

SQ

Query Match 14.2%; Score 211.4; DB 15;
 Best Local Similarity 99.5%; Pred. No. 6.1e-42;
 Matches 212; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 786 CCAAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 845
 DB 1 CCAAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
 OY 846 CTTGGCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 905
 DB 61 CTTGGCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 OY 906 TGGCTCAATGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 965
 DB 121 TGGCTCAATGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 OY 966 GCCAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 998
 DB 181 GCCAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 213

RESULT 10

AA192045 ID AA192045 standard: cDNA: 443 BP.

AC AA192045;

DT 06-NOV-2001 (first entry)

DE Human polynucleotide SEQ ID NO 12105.

KM Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorders; arthritis; inflammation; ss.

OS Homo sapiens.

PN WO200164835-A2

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

PA (HYSE-) HYSEQ, INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI: 2001-514838/56.

DR P-PSDB: AAO12114.

PT Isolated nucleic acids and polypeptides, useful for preventing

PT diagnosing and treating e.g. leukaemia, inflammation and immune

PT disorders

PS Claim 1; SEQ ID NO 12105; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and

CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation. Which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 443 BP; 110 A; 115 C; 129 G; 89 T; 0 other;

SQ

[illegible]

PS Example 2; Column 109-112; 43bp; English.

XX This sequence represents the coding region for the E. coli threonine
CC deaminase enzyme. The gene can be used in a method to produce transgenic
CC soybean plants which have manipulated metabolic pathways that optimize
CC levels of substrates employed in the biosynthesis of copolymers of
CC 3-hydroxybutyrate and 3-hydroxyvalerate.

SO Sequence 1545 BP; 326 A; 437 C; 468 G; 314 T; 0 other;

Query Match 4.7%; Score 69.2; DB 20; Length 1545;
Best Local Similarity 45.5%; Pred. No. 4.2e-07;
Matches 291; Conservative 0; Mismatches 343; Indels 6; Gaps 1;

```
OY 499 CCTCAGGGGGTATGGGGCATGCGTGCCTGATGCTGATAGAGAGCTGGGATTCCTG 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 254 ctctcgagggttaacacacgcgcagggcgctgcttctctctctcgcggttaggcgtgaag 313
OY 559 CCACCATGCTGCTCCCGAGAGACCTCCCTGCAGGTGTGACAGAGCTGCAGGGGAGG 618
    || |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 314 cctgctgcttagtccaaccgcacccgcgcgaatacaagtcgcgcgtcgcggttcg 373
OY 619 GGGCCGAGGTTAGCTGACTGGAAGGTGTGGACGAGGCCAATCTGAGGCCCAAGAGT 678
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 374 gcgcgaaagtgcgtccacgcgcgaactttagaagcgaagcgaagcgtcgaac 433
OY 679 TGGCCAGAGGAGGAGCGCTGGGAAATGTCCTCCCTTTGACCACTCCCTAATATGAAAG 738
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 434 tglcaagaagaagaagggttaccctgggtgcgcgttcgacacacccgaatggtatgcg 493
OY 739 GCCACGCCAGCTGTGTGACGAGACTGAAGAGCTGTGAGGAGCCACCAAGTGCCTCG 798
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 494 ggcgaagcagctgcgcgcgtgaactc-----ctcagaagaagccacatcgcgcgcg 547
OY 799 TCGTGACAGTTGGGGTGGGGGCTCTCTGCGCGGGGTGTGCTGCTGCTGAGAGTGG 858
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 548 tatttgcacatgcgcgcgcgcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 607
OY 859 GCTGGAGAGATGATCCATCATTTGCCATGAGAGACCATGAGGAGACTGCTTCAATGCGG 918
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 608 tgatgcgaagaatcaaaagtgatccgttagaagcgaagacacccgtcgtcgtcgtcgtcgt 667
OY 919 CCATCACAGCCGGAAGCTGTCTACACTTCACAGATCTCCAGATCTGGCCAAAGAGCCTGG 978
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 668 cgcgtgagtcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 727
OY 979 GTGCCAAGAGCGGTGCGCTCGGGCCCTGGAAGTGCATGAGGTGTGCAAGATTCACTCTG 1038
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 728 cgttaaagaacatcgtgtagcgaacaccccttcgttagcgaagagatcgcgaagacatca 787
OY 1039 AAGTGTGAGAGACACCGAGGCTGTGAGGCTGTGAGAGAGCTCCCGATGATGAGAGCTGA 1098
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 788 tcaacgcgtatagcgtatcgtatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 847
OY 1099 TCGTGTGAGAGCTGCTGCTGTGGGGGACGCTTGTAGCAGACCAT 1138
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 848 cgtgtgcgcgaacacctcgtgcgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgt 887
```

RESULT 15
AAZ24939
ID AAZ24939 standard; DNA; 1545 BP.

XX AAZ24939;

DT 21-DEC-1999 (first entry)

DE Mutant E. coli threonine deaminase (L447F) coding sequence.

KM Threonine deaminase; enzyme; transgenic plant; soybean; substrate; ds;
KM metabolic pathway; biosynthesis; 3-hydroxybutyrate; 3-hydroxyvalerate.
XX

OS Escherichia coli.

XX Synthetic.

PN US5959179-A.

PD 28-SEP-1999.

PF 13-MAR-1996; 9605-0614877.

PR 13-MAR-1996; 9605-0614877.

XX (MONS) MONSANTO CO.

XX Connor Ward, Jr., Fedele MJ, Clemente TE, Howe AR, Rozman RJ;

XX Hinchey MAM, Fry JE;

XX WPI; 1999-570778/48.

PS A new method for transforming soybeans

XX Example 2; Column 113-114; 43bp; English.

CC This sequence represents the coding region for the L447F mutant E. coli
CC threonine deaminase enzyme generated by site-directed mutagenesis from
CC the wild type sequence (AAZ24936). The gene can be used in a method to
CC produce transgenic soybean plants which have manipulated metabolic
CC pathways that optimize levels of substrates employed in the biosynthesis
CC of copolymers of 3-hydroxybutyrate and 3-hydroxyvalerate.

SO Sequence 1545 BP; 325 A; 437 C; 468 G; 315 T; 0 other;

Query Match 4.7%; Score 69.2; DB 20; Length 1545;
Best Local Similarity 45.5%; Pred. No. 4.2e-07;
Matches 291; Conservative 0; Mismatches 343; Indels 6; Gaps 1;

```
OY 499 CCTCAGGGGGTATGGGGCATGCGTGCCTGATGCTGATAGAGAGCTGGGATTCCTG 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 254 ctctcgagggttaacacacgcgcagggcgctgcttctctctctcgcggttaggcgtgaag 313
OY 559 CCACCATGCTGCTCCCGAGAGACCTCCCTGCAGGTGTGACAGAGCTGCAGGGGAGG 618
    || |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 314 cctgctgcttagtccaaccgcacccgcgcgaatacaagtcgcgcgtcgcggttcg 373
OY 619 GGGCCGAGGTTAGCTGACTGGAAGGTGTGGACGAGGCCAATCTGAGGGGCAAGAGT 678
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 374 gcgcgaaagtgcgtccacgcgcgaactttagaagcgaagcgaagcgtcgaac 433
OY 679 TGGCCAGAGGAGGAGCGCTGGGAAATGTCCTCCCTTTGACCACTCCCTAATATGAAAG 738
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 434 tglcaagaagaagaagggttaccctgggtgcgcgttcgacacacccgaatggtatgcg 493
OY 739 GCCACGCCAGCTGTGTGACGAGACTGAAGAGCTGTGAGGAGCCACCAAGTGCCTCG 798
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 494 ggcgaagcagctgcgcgcgtgaactc-----ctcagaagaagccacatcgcgcgcg 547
OY 799 TCGTGACAGTTGGGGTGGGGGCTCTCTGCGCGGGGTGTGCTGCTGCTGAGAGTGG 858
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 548 tatttgcacatgcgcgcgcgcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 607
OY 859 GCTGGAGAGATGATCCATCATTTGCCATGAGAGACCATGAGGAGACTGCTTCAATGCGG 918
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 608 tgatgcgaagaatcaaaagtgatccgttagaagcgaagacacccgtcgtcgtcgtcgtcgtcgt 667
OY 919 CCATCACAGCCGGAAGCTGTCTACACTTCACAGATCTCCAGATCTGGCCAAAGAGCCTGG 978
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 668 cgcgtgagtcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 727
OY 979 GTGCCAAGAGCGGTGCGCTCGGGCCCTGGAAGTGCATGAGGTGTGCAAGATTCACTCTG 1038
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 728 cgttaaagaacatcgtgtagcgaacaccccttcgttagcgaagagatcgcgaagacatca 787
OY 1039 AAGTGTGAGAGACACCGAGGCTGTGAGGCTGTGAGAGAGCTCCCGATGATGAGAGCTGA 1098
```

Db 788 tcaccgtagatagcgatgcatctgtcgagatgaagattatctgaagatgtgagcg 847

QY 1099 TGTGTGTGAGACCTGCTGTGTGGGACGCTTAGCAGCCAT 1138

Db 848 cgtgtgaggaaccctctgtgagcgctgtgagcgaggaat 887

Search completed: June 5, 2002, 01:22:15
Job time: 3997 sec


```

OY 786 CCAGTGGCCCTGCTGCTGCAATTGGGGGTGGGGGTCTCTGCGCCGGGATGGCTGGC 845
DB 1 ccaagtgcccttgctgctgcaagtcgggggtggtctccttgccgggtggtgctggc 60
OY 846 CTGCTGAGAGTGGGCTGGCGCAGCATGTATCCCATTTGCCATGAGAGACCATGGGACAC 905
DB 61 cgcggaagtggtgctgcaatgacatcatgagacccatgagacccatgagacac 120
OY 906 TGCTTCAATGCGGCATCAACAGCCGCAAGCTGTACACATTCCAGACATCAACAGTGTG 965
DB 121 tgcctcaatgcggccatcacagcccgcaagctgtacacattccagacatcacagtg 180
OY 966 GCCAAGAGCTGGGTGCAAGAGAGTGGCGCT 998
DB 181 gccaaagacctgggtgccaagacgtggccgct 213

```

RESULT 10

RESULT 9

AA057448 ID AA057448 standard; CDNA to mRNA; 213 BP.

AA057448; AC

19-OCT-1994 (first entry) DT

Ser dehydrogenase like protein. DE

Human CDNA; library; enzyme; protein; ss. KW

Homo sapiens. OS

0920359517 PN

17-FEB-1994. PD

04-AUG-1993; 93WO-JP01095. PF

04-AUG-1992; 92JP-0208077. PR

13-NOV-1992; 92JP-0327619. PR

26-FEB-1993; 93JP-0061431. PR

(SAGA) SAGAMI CHEM RES. CENTRE. PA

Iwahori A, Kato S, Kato T, Kim N, Oh S, Sekine S; PI

WPI: 1994-065688/08. DR

P-PSDB; AAR46109. DR

CDNA of human origin and proteins coded by it - which may be PT expressed by in vivo or in vitro translation using sense RNA or

antisense DNA corresponding to the CDNA. PT

Claim 1; Page 68; 167pp; Japanese. PS

mRNA expressed in human fibrosarcoma cell line HUT-78 was CC isolated and used to construct a CDNA library using vector

CC pTZ18Rpl. Clone HP00157 encoding Ser dehydrogenase-like protein CC was isolated.

Sequence 213 BP; 34 A; 62 C; 77 G; 40 T; 0 other; SQ

Query Match

Best Local Similarity 14.2%; Score 211.4; DB 15; Matches 212; Conservative 99.5%; Pred. No. 6.1e-42; Mismatches 0; Indels 0; Gaps 0;

RESULT 15
 AA573827/c
 LOCUS AA573827 614 bp mRNA linear
 DEFINITION nk08b03.s1 NCI_CGAP_Co2 Homo sapiens cDNA clone IMAGE:1012877
 similar to SW:SDHL_RAT P09367 L-SERINE DEHYDRATASE ; mRNA
 sequence.
 ACCESSION AA573827
 VERSION AA573827.1 GI:2348342
 KEYWORDS EST
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 614)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
 Ph.D.
 cDNA Library Arraying: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 1319 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 445.
 FEATURES
 source Location/Qualifiers
 1..614
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1012877"
 /clone_lib="NCI_CGAP_Co2"
 /tissue_type="tumor"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: colon; Vector: Bluescript SK-; Site_1: EcoRI
 ; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
 dT. Bulk colon villous adenoma. 5' adaptor sequence: 5'
 GAATTCGGCAGAG 3' 3' adaptor sequence: 5'
 CTCGAGTTTTTTTTTTTTTTTTTT 3' Average insert size: 1.1 kb."
 BASE COUNT 126 a 185 c 179 g 124 t
 ORIGIN

Query Match 39.3%; Score 583.6; DB.9; Length 614;
 Best Local Similarity 98.7%; Pred. No. 4.2e-122;
 Matches 610; Conservative 0; Mismatches 4; Indels 4; Gaps 2;

QY 861 TGGCAGCATGTACCCATCATTGCCATGGAGACCCATGGGGCACACTGCTTCAATGCGGCC 920
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 614 TGGCAGCATGTACCCATCATTGCCATGGAGACCCATGGGGCACACTGCTTCAATGCGGCC 555
 QY 921 ATCACAGCCGGCAAGCTGGTCACTTCCAGACATCACCAGTGTGGCCAAGAGCCTGGGT 980
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 554 ATCACAGCCGGCAAGCTGGTCACTTCCAGACATCACCAGTGTGGCCAAGAGCCTGGGT 495

OY	981	GCACAAGACGGTGGCCGCTCGGGCCCTGGAGTGCATGCAGGTGCGAAGATTCACTCTGAA	1040
Db	494	GCCAAAGAGGTGGCCGCTCGGGCCCTGGAGTGCATGCAGGTGCGAAGATTCACTCTGAA	435
OY	1041	GTGTGTGAGGACACCCGAGGCGTGTGAGGCGTGTGCAGCAGTCTCTGATGATGAGCGTATG	1100
Db	434	GTGTGTGAGGACACCCGAGGCGTGTGAGGCGTGTGCAGCAGTCTCTGATGATGAGCGTATG	375
OY	1101	CTGTGTGAGACCTCGCTGTGTGGGACACCTTTAGACGCATCTACTCAGGCGTCTCGGAG	1160
Db	374	CTGTGTGAGACCTCGCTGTGTGGGACACCTTTAGACGCATCTACTCAGGCGTCTCGGAG	318
OY	1161	CTCCAGAGCCGAGGCGTCTGCCCCCTTCCCTGACTTCAGTTGTGTGATGTGTGTGGA	1220
Db	317	CTCCAGAGCCGA-6GCTGCTCGCCCCCTTCCCTGACTTCAGTTGTGTGATGTGTGTGGA	259
OY	1221	GGCAACAACATCAACAGCCGAGAGCTGTGACAGCTTTGAAAACCCACCTGGGGCAGTCTGA	1280
Db	258	GGCAACAACATCAACAGCCGAGAGCTGTGACAGCTTTGAAAACCCACCTGGGGCAGGCTGA	199
OY	1281	GGGGTCCCATCTGGGCCCAAGAGCCCTGAGAGGCCCATGACAGACTCCGTGTTCGAT	1340
Db	139	GGGGTCCCATCTGTGCCCCCAAGAGCCCTGAGAGGCCCATGAGAGTCCGTGTGTGAT	139
OY	1341	GAGAGAGACTCGTCTGGCAGATGGCAGTGGAGAGCTGCGCTGTGCAACTGTGCTGCTG	1400
Db	138	GAGAGAGACTCGTCTGTGCGAGATGGCAGTGGAGAGCTGCGCTGTGCAACTGTGCTGCTG	79
OY	1401	CCGCCGAAGAGAGCCCTCTGAGACTGCTCTTTGGCTCTCCGACAACACTCGGCCAATA	1460
Db	78	CCGCCGAAGAGAGCCCTCTGAGACTGCTCTTTGGCTCTCCGACAACACTCGGCCAATA	19
OY	1461	AACACTTTCGAATTGAA	1478
Db	18	AACACTTTCGAATTGAA	1

Search completed: June 5, 2002, 00:42:48
Job time: 1836 sec

